

Development of a New Assessment System to Evaluate Students' English  
Communicative Capacities in China

Xiaomin Pan

Department of Graduate and Undergraduate  
Studies in Education

Submitted in partial fulfillment  
of the requirements for the degree of  
Master of Education

Faculty of Education, Brock University  
St. Catharines, Ontario

© Xiaomin Pan 2013

## **Abstract**

The purpose of this study was to develop a new English assessment system to evaluate Chinese students' English communicative capacities. Since there is more interaction with people from English-speaking countries, Chinese people attach more importance to English oral skills, and a lot of Western English teaching methods were introduced into China to improve students' English communicative capacities. However, traditional paper-written examinations, like the English test of higher education entry examination, cannot evaluate it effectively. This study explored the perceptions of two Chinese English-language teachers and two Chinese students about English assessment system. A qualitative research method using telephone interviews was conducted in this study. The findings showed that the most possible ways to assess Chinese students' English communicative capacities were paper-written examination and person-machine conversations, although measures should be taken to improve these two models. On the other hand, the model of person-person conversation was the ideal assessment tool but was hard to achieve at the current stage.

## **Acknowledgements**

This Master Research Paper reports on research conducted in the Graduate Studies in Education, Master of Education (M Ed) program at Brock University, St. Catharines, ON. I would like to acknowledge certain individuals who have provided support and encouragement throughout my study.

To Dr. Coral Mitchell: Thank you for your encouragement, patience, and professionalism, which guided me through the project process. Every time I ran across difficulties, you always analyzed the problems with me and gave me professional suggestions to clear my confusion. You are such a brilliant supervisor, and it was my honour to become your student. Your editorial comments are also useful for my further study.

To Dr. Xiaobin Li: Thank you for being the second reader of my major project paper. Your profound knowledge really left me with a deep impression. Our similar backgrounds also made me feel comfortable when I discussed my research project with you. Thank you for all the useful suggestions, and I think your successful career in Canada also set up a model for all the Chinese international students here.

To Heying Yin and Zimin Liu: You are my family and the most important people to me in the world. Although you are far away from me in China, your love and support are always with me. Thank you for listening to me when I complained about the difficulties I ran across during my study. Thank you for your comfort when I felt lonely in a different country. You are the people I would like to share my joy with mostly. Thank you.

## Table of Contents

	Page
Abstract	ii
Acknowledgements	iii
List of Tables	vi
 CHAPTER ONE: INTRODUCTION TO THE STUDY	 1
Background of the Study	1
Context of the Study	6
Purpose of the Study	9
Rationale of the Study	9
 CHAPTER TWO: LITERATURE REVIEW	 12
Old Chinese English Curriculum Requirements	12
Current English Curriculum Standards at Secondary Stage	14
Current English Assessment System	17
Experimental English Assessment Systems	22
Summary	28
 CHAPTER THREE: METHODOLOGY AND PROCEDURES	 29
Site and Participant Selection	29
Data Collection	30
Data Analysis	36
Credibility of Data	38
Ethical Issues	40
Limitations of the Study	41
 CHAPTER FOUR: FINDINGS	 43
Paper-Written Examination	43
Person-Machine Conversation	50
Person-Person Conversation	56
Systemic Changes for the Model of Person-Person Conversation	62
Summary	69
 CHAPTER FIVE: SUMMARY, DISCUSSION, AND IMPLICATIONS	 70
Summary of the Study	70
Discussion	71
Implications	80
Conclusion	84
 References	 85
 Appendix A: Interview Protocol for Teachers	 88



## **List of Tables**

	Page
Table 1: The Learning Requirement of English Levels at Primary and Secondary Stage in China	15

## **CHAPTER ONE: INTRODUCTION TO THE STUDY**

To promote technological progress and financial development, increased numbers of Chinese people pay attention to English as a necessary language tool to communicate with English-speaking countries. At the same time, more and more people attach great importance to English-teaching methods of those countries, which focus on communicative skills. However, it is hard to evaluate the effects of these new teaching methods because of the paper-written examination system, which is mostly used in secondary schools in China. I am undertaking a research project to find out a new assessment system to evaluate English communicative capacities that would be appropriate for the educational system of secondary schools in China. With an effective assessment system, instructors could evaluate the students' English communicative skills effectively and choose the most useful English-teaching approaches to help them improve these skills.

### **Background of the Study**

There are numerous people learning English in China. Actually, "China boasts the largest English learning population in the world, and a history of over six decades of English teaching and learning" (Li, 1984, p. 2). At present, English is the only main foreign language subject that is a part of the 9-year compulsory education in China.

The significant role of English in China helps to explain the importance and urgency to find an effective assessment system for English learning. The reason why Chinese people are keen on English learning has several different aspects. In terms of this point, Pan (2011) said, "in China, English education carries two layers of

significance: that is, for individual development and for social development” (p. 251).

First, for the aspect of the individual, especially for Chinese students, English is one of the most important compulsory subjects from the time they are in primary schools. For Chinese students, “English proficiency is increasingly recognized as a passport to better education, employment and success” (Pan, 2011, p. 259). To improve their chances in this situation, many Chinese students take extra English classes in their free time. In one employment opportunity, Chinese-foreign joint ventures or foreign-only corporations are always the first choice for the graduates because the salary and the benefits are much better than the local companies, but those corporations have very strict requirements for English skills. For those people who would like to find a decent job in a modern city, like Shanghai, they have to speak very good English. In other cases, for those people who would like to study abroad, they have to pass various English proficiency tests, such as The International English Language Testing System (IELTS) and Test of English as a Foreign Language (TOEFL). Hence, large numbers of students enroll in different kinds of language training centers every year, to prepare for those language proficiency tests.

Second, from the aspect of social development, China is keen on improving its social status in the world. On this point, Pan (2011) said,

[T]he states’ relative strengths depend not merely on the degree to which they can effectively exercise authority internally but also on the degree to which they can compete with the other states in the environment of the world-system in order to move to a more central position in the hierarchy. (p. 246)



In an attempt to compete more effectively, English is regarded as a language tool to learn the advanced technology from other English-speaking countries. Both Hu (2005) and Wang and Lam (2009) note that, since 1986, people's focus has been on using English for the goal of economic development and national modernization. Thus, as the key educational institutions, the schools and universities are encouraged to foster students' English ability as much as possible.

As English teaching developed in China, there were remarkable changes of English curriculum with the change of macro-level political context. For example, in the late 1950s, because the relationship between the Chinese government and Western countries was not close, few foreigners went to China. Hence, most of the English-teaching focused on understanding the reading materials written in English. By contrast, at present, with the widespread Open-Door Policy, Chinese people have more and more opportunities to be in touch with English native speakers. As a result, the attention has been turned from basic comprehension to communicative skills.

In China, there are three levels of agencies involved in the English curriculum development: (a) the State Education Commission as the superordinate, (b) linguistics experts and specialists in language instruction as the intermediate, and (c) the classroom teachers as the subordinate agency (Adamson & Morris, 1997).

The superordinate level "is charged with national curriculum policy decisions-most notably the People's Education Press (PEP), which produces national syllabuses, textbooks, and other resources" (Adamson & Morris, 1997, p. 5). During the development of Chinese English curriculum, this agency plays the core role. For

example, there was an unprecedented revision of English textbooks in the early 1990s. Specifically, it was the first time that external bodies, such as Longman and the United Nations Development (UNDP), participated in the development of Chinese English textbooks. This innovation process and the revised English textbooks were both approved by the State Education Commission.

If any changes were to be made to the assessment system of English communicative skills, the State Education Commission would make decisions at the level of policy. For example, the Commission would decide if there should be culture information related with English-speaking countries taught in English classes, if there should be financial support for English teachers' training, if more native English teachers should be hired, and so on. The Commission would decide (a) if the class size should be reduced in order to provide more practice opportunities to the students, (b) if schools of the same province have relatively equal English education resources, or (c) if measures should be taken to avoid the unfair results of the higher education entry examination due to the unequal education resources of different schools.

During English language development, the State Education Commission consults with the second level of agencies. According to Adamson and Morris (1997), "The intermediate group includes linguistics experts and specialists in language instruction, who are based primarily in tertiary institutes. Members of this group at the provincial, country, and municipal levels also are responsible for public assessment" (p. 5). When any revision of English curriculum is required, the State Education Commission asks the experts and specialists to do the related research. For example,

in the late 1950s, “While the PEP dictated the development of the curriculum during this period, the actual writing of the syllabuses and textbooks was placed in the hands of a variety of intermediate agencies” (Adamson & Morris, p. 11).

If a new English assessment system were to be developed, the members of the intermediate agencies would probably do a lot of research work, such as the form of the assessment, the detailed topics of the questions, the training of the teachers, and so on. At the pilot stage of a new assessment system, those English experts could also be the raters of the assessment. In this way, they could check first-hand the teachers’ teaching quality and they could examine the effectiveness and efficiency of the assessment system.

The third level of agencies is where the decisions are put into action. “The subordinate group consists primarily of classroom teachers who directly implement the syllabus” (Adamson & Morris, 1997, p. 5). Although this group is mostly responsible for carrying out the final decision, they do play a role in the change process. For example, English teaching experienced a search for improving the quality of education from 1960 to 1966, and there was a big change to the English pedagogy and textbooks. During this period, “Subordinate agencies also were involved by promoting change (through experiments with new pedagogy or by criticizing current resources and practices) and by testing new textbooks” (Adamson & Morris, p. 15).

To implement a new assessment system of communicative skills, an innovation of English teaching methods is also required in order to make sure that the

students perform well in the tests and exams, which is a big challenge for Chinese English teachers who are used to the traditional English teaching approaches. On the other hand, they have the opportunity to give feedback about the new assessment system to provide useful information for the intermediate agencies to consider.

These three levels of agencies are closely related to each other. Each agency plays a different role in the innovation of the English assessment system. The superordinate agency makes the decision of revision, the intermediate agency provides the professional suggestions, and the subordinate agency conducts the final practice.

Right now, Chinese English education is faced with another innovation, changed from basic comprehension skills to communicative capacities, which requires all three agencies to collaborate with each other.

### **Context of the Study**

Even though Chinese students begin to study English in primary school, and some people start even earlier, they still find it challenging to communicate effectively with native speakers. From my experience working in language training centers in Shanghai, which features the high English level in China, I found that although some of the senior high school students can attain the English proficiency scores the universities require, few of them can communicate with English native speakers as well as they can achieve in the standardized paper-written tests.

One reason for this result is that when senior high graduates pay attention to English syntactic knowledge and comprehensive skills, the human factors of

English-speaking countries, such as culture and history, are ignored both in the process of English teaching and in the higher education entry examination. As a result, the incomprehension of the culture of English-speaking countries becomes a big barrier when they communicate with the people from abroad. Unfortunately, the Chinese government does not appreciate the culture of English-speaking countries or its role in communicative ability. About this point, Pan (2011) said, “Chinese governments were keenly aware of any potential linguistic and cultural threat” (p. 256). They just regard English as a way to develop the economy and technology instead of a culture worth appreciating. This economic and technological purpose, which has been claimed as the teaching objective in the English teaching syllabus over the years, can be traced back to the late Qing Dynasty. “When the study of English was introduced to China, the official guideline for learning defined ‘Chinese knowledge as the foundation and Western knowledge for utility’” (Pan, p. 256). Although in the current syllabus there is also a requirement for the students to know the culture of English-speaking countries, “emphasis is to help learners consolidate their own subjectivity and cultural identity so as to create a desirable societal scenario for cultural governance” (Pan, p. 257).

As a result, “no matter which level the policy regulation is at (primary, secondary or tertiary), Chinese institutional policies do not grant clear privilege to any type or variety of English” (Pan, 2011, p. 254). For example, it is well-known that there is American English and British English, but in China there is no specification of which country’s language model is to be followed. English learners are just

required to speak “correct, natural, appropriate, fluent pronunciation and intonation” (Pan, 2011, p. 254). Thus, it is not surprising to see that a Chinese English learner confuses whether he or she speaks American English or British English.

Human factors are undoubtedly important when people communicate with each other. That could be an essential reason why Chinese students perform poorly when they communicate with native speakers. Human factors are also important in pragmatic assessment. “Some studies have shown that pragmatic expectations and assessments are culture-specific and learners perceive socio-pragmatic elements such as social distance, relative power and status, and the severity of a specific apology situation differently” (Liu, 2007, p. 396). Therefore, it is necessary to include the culture of English-speaking countries not only into English teaching curriculum but the assessment system as well.

Another reason Chinese students are not good at communicative skills is that the traditional English teaching methods are always teacher-centered, and test-taking tricks make up almost all of the class hours. As Rao (1996) said, "Traditionally, English teaching in China is dominated by a teacher-centered, book-centered, grammar-translation method and an emphasis on rote memory" (p. 458). With the focus on knowledge of grammar and vocabulary, students have little opportunity to learn or practice communication skills.

With the innovation of English curriculum in China, there are now requirements about the learning of the culture of English-speaking countries and English communicative skills. However, new problems rear up during the practice.

Secondary schools still use the traditional standardized paper-written examination to evaluate the student's English skills. According to this point, Hudson et al. (1992) said, "the corresponding tests still focus on aspects of linguistic competence such as syntax, vocabulary, and cohesion, to the exclusion of communicative abilities such as pragmatic competence" (as cited in Liu, 2007, p. 391). Hence, currently, in China, there is no proper assessment system to evaluate whether the students' communicative skills are really improved or not. In other words, people cannot tell whether the current English teaching methods in China are really effective.

### **Purpose of the Study**

The purpose of my study was to ask students' and teachers' opinions about how to develop a new assessment system for English communicative capacities in China. I asked them (a) what an effective assessment system would look like, (b) what would help them to implement this kind of assessment system, and (c) what made it difficult for them to implement this kind of assessment system.

### **Rationale of the Study**

The reason I conducted this study is because Western English teaching methods are currently very popular in China. With the dissatisfaction of traditional English education in China, people are introducing the English-teaching resources from English-speaking countries, including linguistic specialists and English-teaching methods. Even "the policy makers in China seem to adhere strictly to a belief that the native speakers of English are the best teachers and English-speaking countries set the standards" (Pan, 2011, p. 255). A lot of language training centers and government

schools introduce the English-teaching methods from Western countries, such as creating a real language environment for English-learners, pragmatic-skills-oriented instead of exams-oriented assessment, student-centered rather than teacher-centered pedagogy, and so on (Li, 1984).

Although there are various kinds of English teaching methods in China right now and most of them boost its great effect to improve Chinese English learners' communicative skills, if there is no assessment system to evaluate, people could not tell which one is really effective or not.

On the other hand, both English learners and English teachers have different responses to the Western English-teaching methods. Some students like this kind of interactive teaching method; they think it is more interesting than just recitation of the vocabulary and grammar points. Some Chinese parents prefer the English classes taught by native speakers because they think native English-speaking teachers are more professional than Chinese English teachers. On the other hand, some people argue that the students taught by the new communicative approach make more errors in their English than those taught according to the traditional method (Li, 1984, p. 12). In terms of different responses from different angles, it is hard to decide if Western English teaching methods are good or not.

While people have various opinions about Western English pedagogies, the majority of Chinese people regard the marks of the final higher education entry examination as a standard to decide whether this teaching method is effective or not. However, more and more people begin to doubt whether standardized paper-written



exams can evaluate English pragmatic skills. As a result, it is important to investigate what kind of assessment system would evaluate effective English communicative capacities.

My study provided opinions from both English learners and English teachers to the agencies of the State Education Commission as the superordinate level for their decision making or other people who are interested in the innovation of English education.

## **CHAPTER TWO: LITERATURE REVIEW**

To prepare for this study, I read many articles about English language education in China. These articles are related to the old and current Chinese English curriculum and traditional and experimental English assessment systems. This literature helps to explain the innovation of Chinese English curriculum and what corresponding changes should be made to the assessment system.

### **Old Chinese English Curriculum Requirements**

In China, the relationship between micro-level politics and the development of English language education is quite close. Adamson and Morris (1997) analyzed Chinese policy toward secondary English curriculum from 1956 to 1997. They note that English curriculum, textbooks, learning requirement, and teaching methods all received changes corresponding to the policies of different periods. The trend of changing the old Chinese English curriculum can be analyzed in three aspects: purpose, textbooks used, and implementation.

With the industrial expansion of the mid-1950s, the government considered English as a valuable language tool to develop economy and technology. According to this point, Adamson and Morris (1997) said, “the study of English is regarded as necessary for acquiring technological expertise and for fostering international trade” (p. 3). At the same time, the subobjectives varied from period to period. For example, in the mid-1950s, the objective of the English curriculum was “motivating students to continue to learn English in the future and providing them with the relevant knowledge, skills, and techniques” (Adamson & Morris, 1997, p. 8). At that stage,

there was no specific requirement for different English skills. However, in 1993, with the further development of the Open-Door Policy raised by Deng Xiaoping, Chinese people had more opportunities to get in touch with foreigners. The guidelines made the corresponding changes, which “stated that the principal aim of the new curriculum is to foster communication, which previously had been a relatively marginal goal” (Adamson & Morris, 1997, p. 22).

With the focus of English education changing in different periods, people also made corresponding changes to the content of the curriculum. In terms of this point, one of the most typical aspects was to change the English textbooks from period to period. For instance, the textbooks used in the mid-1950s “centered around reading passages and grammar exercises... Many of the texts are borrowed directly from Soviet textbooks” (Adamson & Morris, 1997, p. 8). Because of the lack of English-education experience, the ability to write English textbooks was not satisfying. As Adamson and Morris described, “The majority of texts are selected political documents and moralizing stories... The English in the textbooks was not the English of any English-speaking country” (p. 10). On the other hand, with more interaction with English-speaking countries and further requirements for English communicative capacities, there was remarkable innovation about the textbooks. In the period from 1993 to 1997, for example, cultural information about the major English-speaking countries was included. Besides, this was the first time that foreign publishing houses were involved in the compilation of language textbooks.

Implementation of English curriculum is a crucial part of the whole process of

English language education. At the early stage, because the textbooks focused on reading passages and grammar exercises, the English teaching approaches left people with a mechanical impression, “which [was] characterized by an emphasis on reading and writing skills, constant references to the learners’ mother tongue, a focus on grammatical forms, and memorization of grammatical paradigms” (Adamson & Morris, 1997, p. 8). With the increasing importance attached to pragmatic and communicative skills, teachers’ English teaching methods also changed. There was more interaction between teachers and students in class, and teachers paid more attention to encourage students to communicate in English. According to Adamson and Morris, “the guidelines advise[d] teachers to use a variety of teaching strategies to create situations for promoting communicative competence” (p. 22).

The trend of innovation of English curriculum in China shows that the focus of English education changed from syntactic skills to communicative skills. Whereas the traditional model emphasized basic grammar and vocabulary, the current English curriculum in China attaches more importance to being able to speak fluent English.

### **Current English Curriculum Standards at Secondary Stage**

At present, the curriculum has nine English levels for primary and secondary education in China. The first eight levels are part of the regular requirements (see Table 1), and Level 9 is an extension level for specialist schools and able students. The curriculum guide includes overall descriptors for comprehensive language competence for each level. The detailed explanation shows that the requirements for communicative skills at the secondary stage are actually very high. For example, for

Table 1

*The Learning Requirement of English Levels at Primary and Secondary Stage in China*

Stages		Primary Education						Junior			Senior		
								Secondary			Secondary		
Grade	1	2	3	4	5	6	1	2	3	1	2	3	
Levels	No English		1	1	2	2	3	4	5	6	7	8	

Level 8, which the students are supposed to achieve when they graduate from senior high schools, they should “communicate fairly naturally with other English speakers about familiar topics” (Martin, 2005, p. 7). However, the goal of current English education is not always satisfied. In terms of this point, Rao (1996) said, "after studying English for several years, the students might have learned how to analyze sentence structures and how to translate and appreciate English literature, but they remain at a loss when they meet English speakers" (p. 458).

A couple of reasons might account for this poor result. First, there is an obvious conflict between English Curriculum Requirements and the definition of English levels for primary and secondary stages. Pan (2011) argued that in the English Curriculum Requirement, the focus on communicative skills is at the college level instead of primary and secondary level:

As defined in English Curriculum Requirements at Compulsory Education Stage at Senior High Education Stage, the objective of English learning at primary and secondary school is to develop students’ comprehensive language use competence...Later at college level, more emphasis is put on developing students’ listening and speaking competence. (p. 251)

However, in spite of the emphasis in the curriculum on basic language learning, there are also clear requirements for speaking and listening skills for students to reach at each level during their primary and secondary schooling. This causes confusion for both teachers and students.

Second, in the current English Curriculum Requirements, there is no specified explanation of the assessment system that should be used to evaluate the students' English communicative skills. Currently, most of the English assessment in China uses paper-written examinations, which focus on grammar, vocabulary, reading skills, and writing skills. It is questionable whether the current assessment system can evaluate the students' communicative skills effectively, because even if those students do well in paper-written English examinations, they still have serious problems with communicating in English.

### **Current English Assessment System**

*The New English Curriculum for Chinese Primary Schools and Junior/Senior Middle Schools* includes as follows, according to Martin (2005), "In the New Curriculum, assessment should combine formative assessment and summative assessment" (p. 3). For formative assessment, which concentrates on progress during the study process, there are different forms for teachers to evaluate the students' English ability, such as comparison and assessment of classroom learning activities, self-assessment of learning outcomes, a learning portfolio, questionnaires, interviews, feedback from parents, and everyday quizzes and tests (Martin, 2005).

On the other hand, summative assessment, which concentrates on the outcomes at the conclusion of a course, end of term exams, and graduation exams, are the main means of measuring the level of students' comprehensive language competence (Martin, 2005). One of the typical examples of summative assessment is the higher education entry examination system in China. Davey, Lian, and Higgins

(2007) claim that “[t]he importance of the Chinese higher education entry examination is increasing as China undergoes modernization...A university education markedly increases life chances in China, where society, including the job market, is very competitive” (p. 385).

In Chinese families, the higher education entry examination plays an important role. In traditional Chinese thinking, “a high standard of education is associated with social status, ... failure in school is traditionally associated with individual, family and even national shame” (Davey et al., 2007, p. 387). Most Chinese families begin to prepare this entrance examination at their children’s early age. English, as one of three compulsory subjects of the exams, which are Chinese, English, and mathematics, is an essential part of the preparation for the exams.

The English component of the higher education entry examination is different in each province, but the question types are quite similar, including multiple choice, cloze test, filling the blanks, True or False, matching, and writing. These questions are used to evaluate the students’ linguistic competence with English-language syntax, vocabulary, grammar, and comprehension. In recent years, a speaking test is also included in the English component of the entrance examination, but it is only a very small part of the whole exam and the requirement varies from province to province. In some provinces, only those who would like to apply for an English major or who plan to attend English-featured universities are required to take a speaking test. In the other provinces, the marks of the speaking test make up only a small section of the total marks. For example, in the 2011 higher education entry examination in Guang Dong



province, the total marks for the English exam were 150, and the speaking test made up 15 points, which is only 10%.

This situation presents the problem that people, both the teachers and the students, do not pay enough attention in summative evaluation to the speaking test, which focuses on communicative skills. According to the New English Curriculum, formative assessment is as important as summative assessment. The document emphasizes very clearly to “make sure assessment methods are varied and flexible” (Martin, 2005, p. 17). However, if the speaking test does not attract enough attention in the higher education entry examination system, neither the English teachers on the front-line nor the administrators in charge will attach importance to the teaching of communication skills. Obviously, this is a conflict within the requirements of the English levels of the New English Curriculum.

There are also other shortcomings of the higher education entry examination system. As the most important examination in China, the higher education entry examination is undoubtedly the weathervane of the whole student-assessment system. Almost all other regular exams and tests follow its pattern. Luo and Wendel (1999) claim that “the exam has been powerful force in influencing teaching and learning” (p. 65) because the education the students receive during primary and secondary stages is preparing them for the final higher education entry examination. The considerable extent to which people pay attention to the higher education entry examination has bad impacts in many aspects.

First, students become “test-taking” robots rather than independent learners.

Because of the importance of the higher education entry examination, Chinese education is essentially exam-oriented. “Schools prepare students for tests, which sort them for college attendance and beyond. Nothing in the experience of Chinese students encourages them to think critically except in the narrow confines of testing” (Hammond, 2010, p. 3). In order to get a satisfying result in an English exam, for example, the students will try to use different strategies to take the exam instead of trying to improve their English competence. In Song and Cheng’s (2006) research, they concluded that “the high use of inferencing strategies might be typical of Chinese students in learning English” (p. 256). This result is not surprising because those test-taking tricks make up a considerable proportion of English teaching, but they do not have any positive impact on developing a student’s communicative skills.

Second, students have other huge pressures. China has the largest population in the world. For university education, “[t]he number of applicants far exceeds available places; competition is fierce, particularly for entry into prestigious universities” (Davey et al., 2007, p. 385). Besides, people believe that “[p]assing the national college entrance exam makes a great difference in the lives of students. Those who pass the exam bring honor to the whole family” (Luo & Wendel, 1999, p. 63). There is even an old saying that one exam determines one’s life. The pressure on students to succeed in Gao Kao can, to some extent, explain the high suicide rate among Chinese students.

Thirdly, the higher education entry examination system encourages English teaching methods that are rote-learning and examination oriented. Zhang (1995) and

Mi (1998) note that the entrance exam plays a pivotal role in shaping pedagogical practice because developments in secondary education are determined by the higher education entry examination rather than vice versa. In Chinese English classes, teachers spend most of the class hours on intensive reading (in which teachers explain the sentences word by word, and each grammar point appears in the reading materials) because reading questions make up almost one third of the entrance English exam. Additionally, “[t]o promote more students to college, some teachers focus only on the part of the content that is likely to be examined” (Luo & Wendel, 1999, p. 66). In this way, knowledge that is not the focus of the exams, such as communicative skills, will be ignored in English classes.

Finally, English teachers in China also have great pressure. According to Krebs (1996), “Teachers whose students perform well are praised and rewarded. Teachers whose students perform poorly are penalized, and are less likely to be promoted” (as cited in Luo & Wendel, 1999, p. 61). Here, performing well or poorly refers to the students’ performance in the higher education entry examination because it is the standard people use to evaluate a student during his/her primary and secondary education stages. “This has led to a situation in which the aim of school teaching in China is to prepare students for the higher education entry examination rather than to develop their abilities” (Davey et al., 2007, p. 392). In other words, the entrance exam is also being used as a crucial standard to evaluate whether a teacher is successful or not.

Consequently, because of the deep-rooted higher education entry examination

system, both English teachers and the students are proficient on test-taking skills.

### **Experimental English Assessment Systems**

Some experiments are currently being undertaken to blend formative and summative assessments of English language learning. In one case, Van Naerssen and Riggenbach (1987) developed a study to evaluate the effectiveness of the oral communication skills of an English for Science and Technology (EST) program. In the description of this program, Van Naerssen and Riggenbach note that “EST students simply want to communicate their ideas to their colleagues. The ‘consumers’ of their English will be their professional peers, native English-speaking colleagues, or other colleagues who use English as a second or foreign language for specific purpose” (p. 219). To evaluate the effectiveness of this program, the objective of their research was to “(a) check whether or not the primary ‘consumers’ of these students’ English (other scientists) will be able to understand them, and (b) determine what levels are crucial for comprehensibility and acceptability for various purposes” (Van Naerssen & Riggenbach, 1987, p. 219).

Van Naerssen and Riggenbach (1987) describe the form of evaluation as follows. Each of the participants gave two speech examples, including an informal, unplanned speech and another more professional, planned speech. In this way, the evaluators had a clear idea about the participants’ communicative ability in different occasions. The evaluators were made up of native English-speaking scientists and teachers of English, who were either native speakers of English or nonnative speakers whose first language was Mandarin. This practice was designed to evaluate the

participants' professional academic communication skills and general language skills.

The results showed that “the Chinese scientists will, on the average, probably be able to perform satisfactorily in an informal professional setting”, while for the other tasks in the professional setting “make a conference presentation, teach an undergraduate course, and conduct a graduate seminar” (Van Naerssen & Riggenbach, 1987, p. 225); the results were not satisfying. For general language skills, “The overall average was not on the high end of the range of acceptability” (p. 226). The two lowest items of the ranking were pronunciation skills and grammar skills. This result is not surprising. Even with years of intensive training of English grammar knowledge, if the knowledge is never adopted into pragmatic use, it is just the knowledge of textbooks.

Van Naerssen and Riggenbach's (1987) evaluation process shows an English learner's communicative skills more effectively than the traditional paper-written exams. However, there are also limitations. First, it is hard to make this practice widespread in China. For the large population in China, the English teaching facilities are quite limited, which means that each class will be very large. In large-size classes, the most used question type is multiple choice because these questions can be checked by machine. For example, in the English exam of the standardized higher education entry examination of Jiangsu Province in 2011, multiple-choice questions made up 85 points, which was about 70% of the total score. It might be possible to popularize the evaluation system of speaking English by having students tape the speech examples, but there is still the problem of finding enough teachers to evaluate them, not to

mention finding the native English teachers.

Second, the English communicative ability of Chinese English teachers has been questioned for a long time. About this point, Rao (1996) said, “Owing to a lack of English proficiency themselves, some Chinese teachers are concerned about not being able to answer spontaneous questions about the target language, sociolinguistics, or culture as they arise from interactions in the classroom” (p. 467). In other words, one of the negative consequences to adopting Western English teaching methods in China is that Chinese English teachers are not prepared at the current stage. Most of the English teachers have never been abroad, so their knowledge about the culture of English-speaking countries is limited. Furthermore, Van Naerssen and Riggenbach (1987) found that the results attained by native English speakers and nonnative English speakers were different. However, it is impractical to have enough native English speakers to evaluate the speaking skills of the graduates. As a result, before a new assessment system for communicative skills is implemented, we should figure out who should give the standards, what the standards are, and who will rate the students.

A second experimental assessment system for communicative skills has been introduced by Liu (2007). This system, known as multiple-choice discourse completion test (MDCT), consists of “presenting the test-takers with a speech act scenario followed by three response alternatives to choose from” (Liu, 2007, p. 393). MDCT has its own scoring system, with two points for the key, zero for the option that was incorrect in grammar or vocabulary, and one point for distractors, which

means that the option “was both linguistically inaccurate and socially inappropriate” (Liu, 2007, p. 393).

The introduction of MDCT makes up for a deficiency in the assessment system of pragmatic competence because it specializes in English learners’ communicative skills. According to Liu (2007), “MDCT can easily be administered in the classroom and in a paper-and-pencil format” (p. 409), which is suitable for the current summative test system such as higher education entry examination. “In addition, the MDCT is easy to score, since it consists of selected-response items and can be scored by machine” (Liu, 2007, p. 409), which makes it suitable for the large number of students in English classes. However, there are still some drawbacks to this test.

First, it is unclear who should generate the multiple-choice questions. For example, MDCT just evaluated test-takers’ responses where apologies were expected. To generate the questions, a group of 30 Chinese university EFL students were invited to give a sheet of paper with an illustration of an apology in Chinese. Afterwards, professional translators translated them into English. The problem was that due to different cultural backgrounds, people from different countries would apologize at different occasions. However, the culture of Western countries is typically ignored in Chinese English classes, so these Chinese students were likely to be unfamiliar with the different settings in Western countries where they should apologize. As a result, there is a need to consider that both native speakers and nonnative speakers should be involved to generate the questions together.

Second, it is also unclear about the appropriate topics for these questions. For MDCT, the topic is about apology. However, for tests and examinations, there should be a lot of topics that cover both the scenes of daily life and academic usage. The proportion of these two different scenes could be different for different levels of the tests and the exams. For example, at the primary and secondary stages, the communicative tests could be more specific to daily use, while at the tertiary stage, they could be more focused on professional usage according to students' different majors. When it comes to the detailed content of these questions, for professional usage, it would be better to have both the native speakers and nonnative speakers, who work in the specified work places in both China or Western countries, provide questions. For the part of daily use, native speakers and nonnative speakers who live in China or in Western countries could also be asked to provide questions. For example, in a survey of Chinese immigrants to Canada, Li (1990) asked about their motives to study English. He found that, after years of living in Canada, they had very clear ideas about what they used English most for in their daily lives. "For immigrants, their primary motives in learning a second language is associated with their immediate needs: knowing the language and the basic skills required to function in the new society" (Li, 1990, p. 33), such as using postal service, filling out forms, writing resumes, listening to radio, and so on. That is useful information when people generate the topics of the test. It is also useful information for updating the topics with time.

Third, the multiple-choice type of question is not suitable for evaluating



students' speaking skills because a paper-written test cannot check the students' pronunciation. Additionally, although the students are asked to select the correct answer from three options, the answer is already made. The teachers will have no idea what the students' own answers might have been without the prompts of the options; and in daily life, there are no prompts.

To make up these disadvantages of multiple-choice questions, people suggested introducing new question types to evaluate students' English communicative capacities. One of them is short-answer questions, which is also a traditional question type of paper-written examination. Compared with multiple-choice questions, short-answer questions have several advantages. First, unlike multiple-choice questions, students have to answer short-answer questions with their own words (Pressley, Ghatala, Woloshyn, & Pirie, 1990). Second, short-answer questions assess more aspects of students' English-language proficiency than multiple-choice questions (Bensoussan & Kreindler, 1990). Third, short-answer questions only require the students to use limited words to answer the questions, which makes it different from an English writing test.

However, short-answer questions also have disadvantages. The biggest concern is how to check the answers of short-answer questions because they do not have standardized answers. Given the large population in China, multiple-choice questions are the most popular question types because they can be checked by machines. In this situation, C-rater, an automated scoring system, has been developed to ease the rating process of short-answer questions (Carr & Xi, 2010).

### **Summary**

The literature reviewed in this section explains the background of the current English curriculum and assessment system in China and the shortcomings of English teaching in China. The literature also shows the problems with trying to evaluate English communicative capacities through the traditional Chinese English assessment system. This review also raises a number of questions that are not yet answered. For example, what is the most efficient or effective model of English language assessment to evaluate students' English communicative skills? Is it possible to adopt a new assessment system in China? What challenges would be involved? To adopt a new assessment system, what measures could the three agencies take? They are all open-ended questions that I planned to address through qualitative research.

### **CHAPTER THREE: METHODOLOGY AND PROCEDURES**

Because of many conflicts between the Western English-teaching methods focusing on communicative skills and the prevailing paper-written English examination system, I conducted individual interview research. This decision put the study into the qualitative research methodology. For the reason why I used qualitative methods to do the research, Berg (2001) said, “[i]n the case of life-worlds, researchers focus on naturally emerging languages and the meanings individuals assign to experience” (p. 10). I asked both English-language teachers and learners for their opinions about how to improve the current English assessment system. Individual interviews helped me get an idea about different people’s opinions on how to deal with the conflicts evident in the review of the literature.

#### **Site and Participant Selection**

I collected the viewpoints from both Chinese English learners and English teachers. The English learners have experienced both traditional English-teaching methods and Western English-teaching methods, which focused on communicative skills. In terms of the English teachers, I interviewed Chinese English teachers who had teaching experience of both basic English-language knowledge and English communicative skills. They gave suggestions about how English communicative capacity could be assessed in China.

As a result, I used purposeful sampling to select my participants and sites. As Patton (1990) mentioned “the standard used in choosing participants and sites is whether they are information rich (as cited in Creswell, 2010, p. 206). In order to

choose the participants who could provide me the most useful information, I contacted the students and English teachers of an English-language training center where I used to work in China. It is specified in the English language proficiency examination in order to prepare the students for going abroad. For example, they have professional training classes for The International English Language Testing System (IELTS). IELTS evaluates the four different English skills including listening, reading, writing, and speaking. To prepare the students for these four skills, this training center offers different English classes. In addition, they also have special classes to teach foreign culture and daily speaking English to help the students adapt to the life of the new country. Most of the classes are one-on-one, which gives the teachers more opportunities to communicate with the students. Actually, this English-language training center sets a good language circumstance to adopt Western English teaching methods effectively. As a result, the students and the teachers there are familiar with Western English teaching methods.

I chose the participants who were best qualified according to my previous working experience in this training center. I chose 2 students and 2 Chinese English teachers I knew. Some of them were not in that language training center any more. However, both the students and the teachers were interested in the development of English teaching in China and willing to share their opinions with others.

### **Data Collection**

I collected the information from different people about their opinions about an effective model of English-language assessment to evaluate Chinese English learners'

communicative capacities. Creswell (2010) noted, “In qualitative research you pose general, broad question to participants and allow them to share their views relatively unconstrained by your perspective” (p. 212). During the literature review, both the study for EST and MDCT have their own drawbacks to evaluate English communicative skills. I needed more information from appropriate people to develop a new English assessment system.

I conducted interviews to do this research. Creswell’s (2010) said, “A qualitative interview occurs when researchers ask one or more participants general, open-ended questions and record their answers” (p. 217). I could find out more possible answers by using open-ended questions to interview Chinese English learners and teachers.

I used telephone interviews because all my interviewees are in China. About this point, Berg (2001) said, “the primary reason that one might conduct a qualitative telephone interview is to reach a sample population that is in geographically diverse locations” (p. 82). Similar to one-on-one interviews, telephone interviews allow me to conduct interviews with the participants personally. In this way, I could get in touch with the interviewees in private, and get more explicit answers. In addition, at the same time, I could also observe the interviewees by their tone to get further information not expressed by words.

With the permission of the participants, interviews were audio recorded. I used the software named Skype to make a phone call to the participants in China and used the software named Audonote Lite to record the conversation in the computer.

When I developed my interview questions, I began with developing a schedule of questions. In terms of this, Selltitz et al. (1959), Spradley (1979), Patton (1980), and Polit and Hungler (1993) mentioned “researchers begin with a kind of outline, listing all the broad categories they feel may be relevant to their study” (as cited in Berg’s, 2001, p. 74). In my research, there were four categories I planned to explore. They were the English assessment system for communicative skills, the higher education entry examination system, the conflicts between them, and the role each agency (the State Education Commission, linguistic experts, and front-line English teachers) played in the process of innovation.

Then, I developed specific questions, which I asked in my interviews. Berg (2001) demonstrated “researching should develop sets of questions relevant to each of the outlined categories” (p. 74). Besides, the questions were open-ended questions. In terms of the advantages of open-ended questions, Creswell (2010) said, “In qualitative research, you ask open-ended questions so that the participants can best voice their experiences unconstrained by any perspectives of the research or past research findings” (p. 218).

To do a rigorous qualitative interview research, I considered the type and the order of my questions. In Berg’s (2001) book, he mentioned that six types of questions could be involved in an interview. They are (a) demographics, (b) policy-related questions, (c) organizational memberships, (d) friends and family involved in policy work, (e) personality style, and (f) leisure activities. When it came to the order of these questions, Berg said, “The specific ordering (sequencing) ...

depend on the educational and social level of the subjects as well as their ethnic or cultural traits, age, and so forth” (p. 74). To make the process of the interview smooth, I put the demographics questions first, such as years of learning or teaching English. Then, gradually, I went deep into the academic questions about the English assessment system. The questions also included subquestions, which Creswell (2010) calls probes. I would like to “[u]se them to clarify points or to have the interviewee expand on ideas” (Creswell, 2010, p. 221).

The first question was about the English learning or teaching experience of my interviewees. For example: How long have you been studying or teaching English? There were also subquestions to help them explore the question thoroughly. For example: What kind of English classes did you have or teach? What English skills did you learn from these classes? Did these classes help you improve your English communicative skills? Did these classes help your students improve their English communicative skills? How did you know that?

The second question was about the current English assessment system. In the literature review part, we found out the importance of the higher education entry examination system in China. I asked my interviewees’ opinions about this issue. For instance: What do you think about the English examination of the higher education entry examination system? Are there any parts of this examination related to English communicative skills? Do you think the English examination of the higher education entry examination system is an effective way to evaluate the students’ English communicative skills? Why or why not? Are there any conflicts between the

assessment system for English communicative skills and the higher education entry examination system?

Then, I introduced them to some experimental assessment systems people used to evaluate the students' English communicative skills, and asked their opinions about these systems. I introduced them to the study of EST by Van Naerssen and Riggenbach (1987) and MDCT by Liu (2007). The third question was: Do you think speech example or multiple-choice was a good way to assess the students' English communicative capacities? Why or why not?

The former question provided my interviewees with some new ideas to develop a new English assessment system. For the fourth question, I would like my interviewees to give their opinions about this problem. The question was: In your opinion, what was the efficient and effective assessment method to evaluate the students' English communicative skills? Why?

Like other experimental assessment systems, there are also some drawbacks of the assessment system the interviewees put forward. Hence, the fifth question I asked was: Please analyze the shortcomings of this assessment system. Besides, I also would like to find out whether there were any conflicts between this new assessment system and the higher education entry examination system. The question I asked was how to adapt it to the higher education entry examination system.

Before the sixth question, I introduced the responsibility the three agencies had during the innovation of the English assessment system. Then, I asked the interviewees' opinions about what the three agencies should do to adopt the new



English assessment system they just raised. For example: To be an English learner or English teacher, what support you need to adapt to the new English assessment system? What support do the three agencies need to provide?

Finally, I asked my participants whether there were any things they would like to add relevant to this topic.

Each interview lasted for around half an hour. There were seven questions for both the students and the teachers. Most of the questions were the same, while the first and the sixth questions were different from teachers and students. To make the participants feel comfortable and make the conversation smooth, I used the participants' first language to conduct the interviews. I interviewed the Chinese English students and Chinese English teachers in Mandarin, and translated them into English afterwards.

During the interview, I followed the questions one by one, but I was also flexible enough to follow the conversation with the interviewees. Besides, I also took notes as interview protocol (see Appendix A and Appendix B) to write down other supplementary information. Creswell (2010) mentioned "An interview protocol is a form designed by the research that contains instructions for the process of the interview, the questions to be asked, and space to take notes of responses from the interviewee" (p. 225). It is a good habit to record the details during interviews; and it is a good way to remind the interviewers to follow the procedure of the interview as well.

### **Data Analysis**

After the interviews, I began to analyze the raw data. About the purpose of data analysis of qualitative research, Berg (2001) notes, “Insights obtained from qualitative research cannot only add texture to an analysis but also demonstrate meanings and understandings about problems and phenomena that would otherwise be unidentified” (p. 102). In the process of data analysis, I better understood the different models of English examination, and further explored whether anything could be done to improve the current English assessment system in order to make it more effective and efficient to evaluate Chinese students’ English communicative capacities.

The way used to analyze qualitative data is content analysis. Berg (2001) explains that “One cannot pull out numbers (operationally reduce responses) from the interviews and expect to plug them into a qualitative analysis computer program-none exists” (p. 102). Although some software is available that could simplify the data organizing process, my study had four interviewees; therefore, I analyzed the contents of the data by hand.

About the detailed process of data analysis, Creswell (2010) sets out six commonly used steps, which are:

preparing and organizing the data for analysis; engaging in an initial exploration of the data through the process of coding it; using the codes to develop a more general picture of the data-descriptions and themes; representing the findings through narratives and visuals; making an

interpretation of the meaning of the results by reflecting personally in the impact of the findings and on the literature that might inform the findings; and finally, conducting strategies to validate the accuracy of the findings. (p. 237)

I used these six steps to organize, code, and analyze the data collected from the participants.

When all of the interviews had been finished, I transcribed all the audio materials into text data. At this stage, I transcribed it first in Chinese because this was the language by which I conducted the interviews with both the teacher and the student participants. Then I translated all the transcripts from Chinese into English by myself.

When all the raw data had been translated, I first organized the data into two different categories, which were Teacher and Student. On each page, I left some margins in which I could make notes and write emerging codes and themes. After the data had been presented in the two categories, I read through the data to obtain a general sense of the material. I noted, for example, what a particular teacher or student participant said about the various test models or what suggestions were made for the current system of the higher education entry examination.

After that, I located text segments that were similar to one another and assigned different code labels to them. For instance, I used yellow color to mark the participants' English-language learning or studying experience. I marked opinions about the strengths of the paper-written examination as dark-green and the weaknesses as light-green. Both teacher and student participants' opinions about the

advantages of the model of person-machine conversation were marked as purple, and the disadvantages of it were marked as orange. The advantages of the model of person-person conversation were marked as dark blue, and the disadvantages of it were marked as light blue. The suggestions from the participants were marked as red. In this way, it was easy to locate different contents after the data had been summarized and analyzed.

To further develop the data-descriptions and themes, I summarized the different codes and developed four themes that described the content within each code. These themes were the teacher and the student participants' opinions about (a) the paper-written examination, (b) the model of person-machine conversation, (c) the model of person-person conversation, and (d) the systemic changes for person-person conversation. I drew meanings from these data-descriptions and themes to develop a new English assessment system for communicative capacities in China.

### **Credibility of Data**

During the research, I had to make sure my findings and interpretations were accurate. Creswell (2010) claimed, "Validating findings means that the researcher determines the accuracy or credibility of the findings through strategies such as member checking or triangulation" (p. 259). Since I used to teach in China, I have my own opinions about the problem of the current English assessment system. I had to stick to my role as an objective researcher rather than interpreting the findings in accordance with my own will. According to Creswell, there are three ways people always use to make sure the accuracy of the findings. They are "triangulation,

member checking, and auditing” (p. 259).

In terms of triangulation, I tried to hear voices from different aspects. Creswell (2010) said, “Triangulation is the process of corroborating evidence from different individuals” (p. 259). Hence, I collected the data from different resources.

Specifically, my interviewees included both English students and English teachers. In this way, I could hear about the issue from different angles.

With regard to member checking, I made sure my participants agreed with the accuracy of the findings of my report. Creswell (2010) noted that “Member checking is a process in which the researcher asks one or more participants in the study to check the accuracy of the account” (p. 259). When I finished transcribing the audio materials into text data, I sent the transcription to each participant and asked them to check the accuracy in order to avoid any misunderstanding during the interview and mistakes of the transcription. In addition, I also sent them the summary of what I obtained from their data so that they could see how I interpreted what they had said. They were asked to check that I had interpreted them correctly or to suggest changes to my interpretations.

When it came to external audit, I needed a second person outside the study to review my report to give objective opinions. About this point, Creswell (2010) recommended that this person “conduct a thorough review of the study and report back, in writing, the strengths and weaknesses of the project” (p. 260). In this way, I could have more nonaligned opinions about my study. Actually, my advisor and my second reader of my major research paper served as the external auditors.

### **Ethical Issues**

My research was approved by the Ethics Research Board of Brock University, and the File Number is 11-259-MITCHELL. I conducted my research according to the guidelines of the Board. Before the interview, I had sent an invitation letter to the people I interviewed in order to introduce them to the background information in detail and ask them for permission to do the interview. When I got the permission to do the interview, I also asked the interviewees to complete an informed consent form to get their consent and clear the confusion about the procedure about the interview. After the interview, I wrote an appreciation letter to each interviewee to thank them for participating in this study and provided the findings of the study to the participants who were interested.

I protected the participants' confidentiality. Creswell (2010) mentioned "Participant confidentiality is of utmost importance" (p. 232). Hence, throughout my study, confidentiality of the participants was ensured. Specifically, I was the only person who had access to the raw data and original transcripts, and I conducted the translation by myself. I did not use any names of participants in any discussions or reports about the project. When I used the data from a participant, I referred to him or her by a pseudonym or number.

I was careful about sharing my own English learning and teaching experience during the interview because it was possible that I could put a positive or negative influence on my interviewees. As an investigator, I was not supposed to bring my own feeling and emotion to the research. About this point, Patton (2002) said, "it may be

necessary to clearly define your role as a researcher (not a therapist offering advice or a judge evaluating the circumstances)” (as cited in Creswell, 2010, p. 231). As an interviewer, I paid attention not to use my personal perception to influence participants’ opinions. Besides, during the interview, I also used neutral words to avoid putting my own viewpoints on the participants and mislead them.

### **Limitations of the Study**

There are also some limitations about my study, which people could work on for further research in the future. First, I chose my Chinese participants based on my working experience in one English-language training center. No matter whether they are still in that institution, they have some features in common, which limits the data resource of my study. For example, this training center features its training of communicative skills so both the teacher and the students are familiar with the Western English teaching methods. If I had chosen my interviewees in different educational organization, the results could have been different.

Second, I know these participants personally. This could lead to the close relationship between the investigators and interviewees, which could cause the results of the study to be inaccurate. To avoid this, prior to the interview, I clearly defined my role as a researcher in the informed consent letter and asked the interviewees to provide their own opinions about the English assessment system for communicative capacities.

Third, I am the only person who has access to the raw data and can analyze them. During the process of interpretation, it is possible that the results could be

affected by my own opinions. Thus, the external audit helped me review the project from different angles and gave me some suggestions I did not think about.



## **CHAPTER FOUR: FINDINGS**

In this chapter, I present the findings of a study undertaken to explore the best way to assess Chinese students' English communicative capacities. Results were drawn from interviews with 2 Chinese English teachers and 2 Chinese English learners. They shared their English teaching/learning experience, and gave their opinions about different examination models to assess Chinese students' English communicative capacities. Data from the interviews were categorized into four themes: (a) paper-written examination, (b) person-machine conversation, (c) person-person conversation, and (d) systemic changes for the examination model of person-person conversation. Each theme includes the participants' opinions about the strengths and weaknesses of the exam model, and their suggestions for further use and development of the model. A summary of the findings is provided at the end of the chapter.

### **Paper-Written Examination**

The paper-written examination is the most popular assessment model used in Chinese schools, from primary schools to universities. As the most typical example of paper-written examination, the higher education entry examination asks students to write the answers to most of the questions on paper. To evaluate this model, I asked participants for their opinions about the higher education entry examination, including their experience of the examination, its advantages and disadvantages, whether they thought the English test of the higher education entry examination could assess Chinese students' English communicative capacities, and any suggestions to improve

the examination. The participants' suggestions were invited in the hope of making this traditional examination model more suitable to evaluate Chinese students' English communicative capacities.

### **Strengths of Paper-Written Examination**

Each of the participants admitted the importance of the paper-written examination. As one of the student participants said, "I think it accords to our national conditions" (H, student). The participants gave several advantages about paper-written examination.

First, a paper-written examination could cover more content, including syntax knowledge, and vocabulary. One of the student participants took the higher education entry examination as an example. She said, "Except the oral test, it involves listening, reading, and writing, what the students learned in their school lives. It can reflect what the students learned in these years objectively, including the students' learning ability and their intellectual capacity" (H, student). When I introduced my participants to the MDCT (Liu, 2007), which is a new type of paper-written examination with multiple-choice questions, Q said,

For the students, if there are more multiple-choice questions, the contents covered in the test will be more comprehensive. In other words, the students at least know the vocabulary of different occasions. Besides, they know which sentence structures they are supposed to use in these situations. (teacher)

In their opinion, a paper-written examination, including the type with multiple-choice questions, could cover more content than an oral test.

Second, since paper-written examination covers more content, it helps improve students' English level. For this aspect, R said,

By preparing for the higher education entry examination, the students could have a better knowledge of grammar. For their further study of English language, including oral English, I think the students will have a better basic knowledge of English language. (teacher)

### **Weaknesses of Paper-Written Examination**

For the negative side of paper-written examination, none of the participants thought that it could assess the students' English communicative capacities effectively. I asked both the teacher and the student participants whether the English test of the higher education entry examination, as the most typical example of paper-written examination, could assess a student's oral communicative capacities. They all had a negative opinion to some extent. S said, "I think it is in an indirect way instead of a direct way" (student).

They also mentioned that some parts of the English test of the higher education entry examination could have some relationship with oral English, but even those parts would not assess a student's English communicative capacities effectively. H said,

Talking about English communicative capacities, there are English listening test and English writing test in the higher education entry examination. I don't think it can do it in a direct way, because it's not a fact-to-face process. Just... it cannot. (Laughing...;student).

This result demonstrates that both the teacher and student participants doubted the effectiveness of the paper-written model of the English test for assessing students' English communicative capacities directly.

Second, participants believed that the paper-written examination makes the teachers concentrate on the exam and ignore the practical English communicative capacities. For example, Q said,

Since everyone's final goal is to go to a good university... the teachers must hope their students get great grades. As a result, they just focus on the exam itself. I mean the content of the examination papers. Comparing with it, I think little knowledge about the English communicative capacities is taught in class.  
(teacher)

Another teacher participant had a similar idea. R said,

I think it is quite mechanized. I think this exam is to evaluate the students' capacity of memory. Generally speaking, a student who works harder or has better capacity of memory can get better marks... I think the English exam of higher education entry examination is exam-oriented. It does not attach any importance to the ability of English practical ability. (teacher)

Both of these excerpts show that the teacher participants were concerned about the pressure the Chinese English teachers have. If the English teachers think the oral part is not important in the higher education entry examination, it is possible that they would skip this part in their classes. About this point, S (student) shared his story:

There were some English oral classes when I was in junior high school. At

that time, my English teacher paid attention to oral English. She thought the practical capacity is much more important than the grades of the higher education entry examination. Oral English is more helpful for the students than just passing the exams. However, when we were in the third grade of the junior high school, she focused on the entrance examination of the senior high school because of the pressure... However, when we had the senior high school entrance examination, we still couldn't compete with other students.

(student)

This story reflects the reality of English education in most of the secondary schools in China. Even if the students' oral capacity is better than other students, they might still fail in the written examination of the English test, which has a negative impact on their ability to gain admission to a good university. Owing to this pressure, teachers make a compromise even if they understand the importance of English oral communicative capacities.

Third, according to the participants, a paper-written examination does not help the students improve their English communicative capacities. For this point, Q said,

Actually, a lot of students' English-language capacities are still very limited after they had the higher education entry examination, especially their communicative capacities. Both their English oral abilities and writing abilities are very poor. All I can say is they made a progress in the aspects of basic English grammar knowledge and English vocabulary. However, it is still at the stage of cognition instead of practical use. (teacher)

The students felt the same way. One of the participants shared that the way to get good marks in the English test of the higher education entry examination was to do the exercises and memorize the answers during a short period.

In sharing his English-learning experience, S (student) said that he got very poor grades in the practice test of the higher education entry examination but that he earned very good grades in the real test. When I asked him how he could make such progress in a short period, he told me it was quite simple.

There was a book named Daily Practice when I was in the senior high school. I copied the book carefully. Next, I did all of the exercises on the book again. Then, I checked the answers thoroughly.... When I saw the questions, I knew which option of the multiple-choice questions was correct. Besides, I knew all of the traps. (student)

Even though he could answer the questions correctly, he did not master the knowledge of those questions. He just memorized the questions and wrote down the answers mechanically. This participant also admitted that even though he received good marks in the exam, he still could not communicate with others in English.

### **Suggestions for Paper-Written Examination**

Given the widespread use of paper-written examinations in China, it is hard to change the current situation in a short time. I asked the participants for their suggestions on how to improve this traditional examination model to more effectively assess Chinese students' English communicative capacities.

H (student) suggested including short-answer questions to assess students'

English communicative capacities, because short-answer questions require the students to give their own answers instead of just selecting the ready-made answers of multiple-choice questions. She said,

I'm thinking about whether there is another way, like short-answer questions. For example, you raise any questions of English oral test... For example, if there's a dialogue between two people, the questions are written on the paper, and you answer the questions by writing instead of speaking. Just in some simple sentences. (student)

S suggested other types of questions that could be introduced into the oral test: There could be many different types of questions combined with oral English and also mixed with the listening part and reading part. In other words... I am not familiar with how many contents need to be covered in the exam. You know what I mean. The contents need to be covered to assess the students' English knowledge. I don't know how many contents need to be covered. If so many contents need to be covered, I suggest there should be a reading part. (student)

According to the student participants' opinions, given the huge population in China, it was more practical to have the paper-written examination to assess the students' English communicative capacities. However, it was necessary to modify the question types of the current English test.

Although they were more likely to use a paper-written examination in China, the teacher participants still worried that the paper-written examination would still

make the teaching exam-oriented, even if the question types were different. Q claimed,

I think it will turn to be exam-oriented in the end. I mean the model of simple-answer questions could evaluate the students' ways of thinking.

However, I'm afraid people will make it exam-oriented. There will be some templates in the future. The answers could be prepared in advance, and it cannot represent the students' real English capacities. (teacher)

Paper-written examination is the most popular examination model in China.

Although each of the participants agreed that it was not the most effective way to assess students' English communicative capacities, they all noted that it was the most practical way of assessment, especially when the exam is widespread in a large country like China. To better evaluate Chinese students' English communicative capacities, the student participants suggested ways to modify the question type of the current English test.

### **Person-Machine Conversation**

The model of person-machine conversation refers to the dialogue between a student and a computer. The question type could be various. The examinee could be asked to listen to a recording and answer the questions, or to have a presentation for 2 or 3 minutes on a certain topic. The examinee's performance would be recorded by the computer and assessed by English teachers afterwards. Since people in some advanced areas use the person-machine conversation for the English oral test in the higher education entry examination, there have been a lot of different opinions about



this new test model. I asked the participants' opinions about this examination model to find out whether they thought it was a good way to assess Chinese students' English communicative capacities. They analyzed the model of person-machine conversation from different angles.

### **Strengths of Person-Machine Conversation**

One of the teacher participants thought the first strength was that person-machine conversation could give the students opportunities to speak English instead of just writing English on papers. Q said, "The advantage of the first model [person-machine conversation] is the students have an opportunity to speak in English. In other words, they could express what they learned by oral" (teacher). She described it as a new examination type that is different from the traditional paper-written examination.

Second, most of the participants thought it was a good way to evaluate the students' English communicative capacities. For this point, H said, "Person-Machine conversation could be more direct, because it could assess a student's language expression ability more directly and more effectively, like English oral capacity" (student). The English teachers also agreed with it. R claimed,

I think the exam model of person-machine conversation is more vivid because it is an actual conversation. You could hear the students' pronunciation and tone, including their usage of vocabulary, grammar, specification of language, and fluency. You could hear all of them directly, so I think it is more vivid.

(teacher)

To their mind, they thought it was a more direct way of assessing oral English communicative capacity than the paper-written examination.

Third, in most of the participants' opinions, students are more comfortable generating their own answers by person-machine conversation. Comparing with the MDCT model of multiple-choice questions (Liu, 2007), they thought person-machine conversation was much more flexible. From the teachers' angle, R noted,

I think person-machine conversation is much flexible than the latter (multiple-choice questions). In other words, it is more flexible to handle the exams. Even in the aspect of marking, I mean it is much more flexible in many aspects, which is easier for both the teachers and the students to handle it.  
(teacher)

In terms of this point, other participants also mentioned that for multiple-choice questions, answers had already been given, so the students were not free to write their own answers.

### **Weaknesses of Person-Machine Conversation**

Participants also saw several drawbacks about person-machine conversation. First, they noted a time issue would limit the content covered in the test. In terms of this, Q described,

The disadvantage is there is a time limitation. They cannot talk about whatever they want to say. There are always some topics that need to be covered, which is quite limited if you want to evaluate the students' overall English communicative capacities. (teacher)

The student participants also had concerns, but the limitation they referred to was different from what the teachers observed. One of the participants mentioned that the sentence structure the students used in their oral conversation could be limited. S said, “Normally, when a person speaks to others, he only talks about what he knows. In other words, he only uses the sentence structure he knows. In this way, it is difficult for you to evaluate his real English level” (student). In other words, if examinees only use the sentence patterns and vocabulary that they know, the test cannot assess the vocabulary and sentences they do not know.

Second, some participants thought talking to a machine makes the English oral test difficult. Q described,

I think it is not user-friendly to ask the students to talk to a machine... because there is a big difference between person-machine conversation and person-person conversation... Speaking with a real person makes you feel like you really communicate with a foreigner. For example, there is facial expression or some body language. It's easy for people to express their thought. However, when it is a person-machine conversation, it is hard to do that like what I just said. Yes. It is impossible for a machine to imitate a real-world scene. (teacher)

Third, some participants mentioned that because there were no standardized answers for the person-machine conversation, examiners' subjective opinions could affect the marks of the students' performance. Q said, “After the test, the recording shouldn't be marked by the examiner him or herself, because it is one-sided or

subjective to some extent” (teacher). For this point, both the teacher and the student participants expressed their worries about the marking standards. Given the Chinese English teachers’ various English levels, they were concerned about how to guarantee the equity of the marking system.

Fourth, the teacher participants thought the model of person-machine conversation could not arouse enough attention to English communicative capacities. Currently, the universities refer to the students’ English oral grades of person-machine conversation only when they apply for university majors that are related to English language. As a result, those students who apply for other majors, which are the majority of the senior high graduates, will not pay much attention to the English oral test.

About this point, R said, “it is not counted in the total scores. It is just one of the references for university admission” (teacher). Q also said,

The grades of the oral test could be ignored. For those students who would like to apply for the majors that are related to English language, it matters a little bit. However, it doesn’t play a very important role, does it? I think it is not important at all. (teacher)

These observations explained the reason the majority of Chinese people do not attach importance to English communicative capacities.

Fifth, the student participants expressed their concerns about the possibility to implement this model in mainland China. About this point, H claimed,

I think it is much more possible to apply the model of person-machine

conversation in more advanced areas in China. In some places in Western China, or in poor areas, or remote areas, I think it is impossible to implement it.

I think this is the limitation. (student)

Even some advanced areas in China still lack necessary equipment to implement person-machine conversation. R mentioned, “We still cannot arrange a computer lab for each examinee. There are around 40 to 50 students in one computer lab. They talk to the computers at the same time. Thus, I think we still don’t have enough resources” (teacher). If availability of equipment is limited in the advanced areas, and more so in less developed areas, the model cannot be equitably implemented.

### **Suggestions for Person-Machine Conversation**

To avoid the possibility that some examiners could be too subjective, Q (teacher) suggested there should be more than one teacher to assess the records of the conversation. She said,

This recording should be marked by a lot of examiners. It is to evaluate the students’ capacity of communicating with a real person. In addition, to ensure the fairness of the test, many examiners should take turns to evaluate the recordings. (teacher).

Since universities refer to the marks of the English oral test only when students apply for English-related majors, participants also suggested that person-machine conversation should be administered to all students in order to have the attention of enough people. About this point, Q said,

It is too limited if it is just implemented for those students who apply for the majors related to English language. As a result, the students will not pay much attention to it. Actually, the aim of the higher education entry examination is to improve the students' English capacities. If it is just implemented for the students of some group, I think it is hard to achieve this goal. Only when every student is required to have the English oral test and this test becomes more important, will the students focus on this aspect. Then, people will practice their English communicative capacities more often in their daily lives.

(teacher)

The model of person-machine conversation has already been used in the higher education entry examination in some advanced areas of China. It is totally different from the traditional model, paper-written examination, and it is not yet widespread in the whole nation. According to the participants' suggestions, huge amounts of money and systemic changes are needed before this model can be implemented, which at the current time can only occur in some advanced areas of the country.

### **Person-Person Conversation**

The model of person-person conversation refers to the oral test that is between an examiner and an examinee. During the interviews, some participants expressed that their ideal model of English oral test was person-person conversation. In terms of the detailed process of this model, R (teacher) gave me the example of the oral test of IELTS, which was person-person conversation. He described,

The oral test of IELTS is divided into three parts. The first part is kind of the introduction part. It doesn't take a long time and the topics are quite simple. In the second part, the examinees are given a topic and some related questions, which the examinees could follow up. The examinees make a speech according to the topic, which lasts for around two minutes. That is the second part. The third part is related to the second part, but it carries out in a deep-going way and the questions are more abstract.

During the interviews, teacher and student participants described why they thought person-person conversation was the ideal model, and it was hard to achieve this model.

### **Strengths of Person-Person Conversation**

The participants identified several strengths about this model. First, it was believed to be a more direct way to assess a student's English communicative capacities. About this point, R said,

According to my experience, as an English-language learner, person-person communication means you can talk to others face-to-face. In other words, if you run across barriers when you communicate with others, there are other ways that could help you understand others, like gestures, face expression, and other factors related with language. I mean when the students cannot understand the examiners, they could look for 'help'. However, if it is in the situation of person-machine conversation, the examinees are totally isolated from other nonverbal factors, which will put them under great pressure.

(teacher)

Compared to both paper-written examination and person-machine conversation, talking to a real person was seen by the participants as more vivid and more direct.

Second, participants noted that person-person conversation makes it possible for examiners to assess the students' performance right after the oral test, which is helpful for the students to improve their English communicative capacities. According to this aspect, Q wondered, "Why don't we evaluate the students' performance right after the oral test? In this way, the students could remember what they said in the test...I think it will be very effective to improve the students' oral capacity" (teacher). This point is important because, in addition to assessing the students' English-language levels, one of the aims of the examination is to help them improve their practical English capacities.

### **Weaknesses of Person-Person Conversation**

The test model of person-person conversation is really an ideal model in a lot of people's opinions, but participants believed that it would be hard to achieve in reality.

First, participants identified two problems related to the teachers. One is the lack of English teachers who could serve as the examiners to assess the students' English communicative capacities. S claimed,

Do you have any ideas how many students have the higher education entry examination every year? If my memory is correct, there are around 6 million



to 8 million students. If the oral test lasts for 20 minutes for each student on average, which is quite short for an oral test, three students need 1 hour, and 6 million students need 2 million hours. If one examiner works for 8 hours a day, we need 450,000 days in total. If one examiner works for 10 days, we need 40,000 to 50,000 teachers totally. Don't you think the human cost is huge?

(student)

Even with enough teachers, a second problem is to make sure that every teacher is qualified to become the examiner of the English oral test. R claimed,

I find the proficiency of Chinese English-language teachers varies from person to person, which is also ingrained. Yes, let us leave those developed cities aside. In those undeveloped and medium-sized cities, the proficiency of the English teachers is worrying. As a result, you could see there are a lot of unqualified English teachers. For example, math teachers teach English classes or the teachers of other subjects also teach English in some remote places in China. Thus, I think it is an ingrained situation. I also think the proficiency of the English teachers is a very important problem. (teacher)

R's concern was that English teachers' own oral English might not be good enough to assess students' English communicative capacities. One of the reasons for his concern is that most of the Chinese English teachers learned their English skills under the traditional English-language education in China, and they might never have taken an English oral test themselves. In this situation, the participants' worry about the oral English level of the examiners of person-person communication is valid.

An issue related to human costs is the problem with money. To implement the test model of person-person conversation not only requires a huge number of qualified teachers but also much equipment. S (student) mentioned that the English teachers in remote areas might have very poor English pronunciation. In this situation, the teachers need tape recorders and computers to help them teach. In some English-language training centers, for example, the teachers use computer software as the supplementary teaching materials to show the students the difference between an American accent and a British accent. This software also costs money.

The money problem leads to a problem over the inequity of education. As the standardized examination, the higher education entry examination is different from IELTS and TOEFL. It is a more widespread examination, and the number of examinees is far greater than the participants of IELTS and TOEFL. In general, the students who take the IELTS or TOEFL test plan to go abroad for further study, and their families have a certain economic foundation to support them to take the IELTS or TOEFL training, which is quite expensive. However, the higher education entry examination is a national examination, and it does not have so much economic limitation. In this situation, the students from the wealthy families could have advanced equipment to help them study oral English. They could even pay for the most professional instructions to improve their oral English. At the same time, the students from the poor areas might not even have a qualified English teacher. About the inequity of the English oral test, S said,

I saw a lot of children in remote areas of China, especially the children in the

mountainous areas. For them, the most advanced equipment when they have English classes is the tape recorder. It is hard for me to imagine what their oral English is like. If you want to teach oral English, some equipment is necessary. However, it is impossible to popularize the advanced equipment in some mountainous areas. Some people say we just implement the oral test in advanced areas. If we do that, how could we keep our basic education fair?

(student)

The gap of wealth is the first problem of inequity identified by the participants.

The second problem related to inequity is the participants' concerns over bribes. Unlike the multiple-choice questions, which have fixed and standardized answers for each question, the answers for oral English tests are flexible and varied from person to person. In this situation, the examiners handle the marking standards totally. About this situation, Q said, "There could be the problem of cheating. For example, some well-informed parents will go to the examiners who will exam their children and offer a bribe. For me, I think it is very possible" (teacher).

All of these problems, teachers, money, and inequity, are the obstacles facing the implementation of the model of person-person conversation in English oral tests.

### **Suggestions for Person-Person Conversation**

To implement the ideal model of person-person conversation, I asked the participants for any suggestions to make it real. However, they could not come up with any practical ways to solve the problems. The suggestions they mentioned were systemic changes, which they admitted would be hard to achieve at the current stage.

Their opinions on this question are presented in the next section.

### **Systemic Changes for the Model of Person-Person Conversation**

In the interviews, I asked the participants about the role they thought the three levels of educational agencies could play in the process of transforming the English assessment system. The three levels are the State Educational Commission, the linguistic specialists and experts in language institutions, and front-line English teachers. Both the teacher and student participants had opinions about what each agency should do to achieve the ideal model of person-person conversation.

### **Suggestions to the State Educational Commission**

For the State Educational Commission, participants thought that as the decision makers, they should issue the orders about the revolution of English assessment system at the policy level. First, the State Educational Commission should issue the order to include an oral English test into the higher education entry examination. In China, the revolution is from the top down. For the process of the revolution, S described, “The State Educational Commission decrees the regulation that the oral test is mandatory in the higher education entry examination next year. Then, those experts will set the proposition and the teachers will teach the related contents” (student).

All of the participants agreed that the higher education entry examination in China is exam-oriented. To increase the students’ admission rates, the teachers only focus on the contents of the exam. Since the English oral test is not mandatory right now, most of the English teachers, students, and parents ignore the importance of

English communicative capacities. About this point, R said,

The State Education Commission is like a baton, and all of the teachers follow the direction it points. For example, if what the teacher teaches is not on the exam, the students will also complain that the teachers are wasting their time.  
(teacher)

According to the participants, only when the State Educational Commission counts the scores of the oral test into the total scores of the English test of the higher education entry examination, will people pay attention to it. Only then will English teachers attach great importance to their pedagogy of oral English, and the students will work hard to improve their English communicative capacities.

Second, the State Educational Commission should allocate money to support the revolution of the English assessment system. The participants discussed many challenges such as the human cost, teaching training, and multimedia equipment as supplementary teaching materials. For all of these problems, economic support is needed from the State Educational Commission. Participants also talked about the problem of inequity, which results from the difference between the rich and the poor. They thought that if the State Educational Commission allocates money to support the English oral education in poor areas (e.g., through teachers' training and purchasing equipment), it would make a big difference in the students' English communicative capacities. On this aspect, Q said,

In terms of the expense, just like what I mentioned, I think it should be the responsibility of the State Educational Commission, right? In other words, if

you want to reduce the students' burden, the country must pay more money.

(teacher)

Third, participants observed that the State Educational Commission needed to issue a policy to stop the problem of bribes. Because of the importance of the higher education entry examination, all of the parents try their best to help their children get high scores. As an improper mean, participants noted that bribes have been used in some exams already. To ensure the equity of the English oral test of the higher education entry examination, the State Educational Commission needs to issue strict policies against bribery. About this point, Q said, "Talking about the problem of bribes, I think the State Educational Commission should punish this behavior. Of course, the schools should punish this behavior first. I mean, for this problem, the State Commission should coordinate with the local institutions" (teacher).

### **Suggestions to the Linguistic Specialists and the Experts of Language**

#### **Institutions**

For the second level, the linguistic specialists and experts of language institutions, participants thought they should be the institutions responsible for providing professional suggestions to the State Education Commission and for making the plans about detailed measures about the revolution.

First, participants thought that the linguistic specialist and the experts research and should develop the new model of the English oral test. They noted that if the State Educational Commission decided to use the model of person-person conversation to assess the students' English communicative capacities in the university examination,

the question types should be totally different from the existing English test. As a professional educational institution, the linguistic specialists and the experts of the language institutions have the authority and the responsibility to formulate the English oral test. About this point, Q said,

In terms of the linguistics experts and specialists in language institutions, they are also very important. Since they are experts and specialists, they could be the first-line teachers. After years of teaching, they become the experts and specialists, right? Thus, they must have a lot of teaching experience. When they are involved into this revolution, they would provide the plans of implementation in details. For example, what kind of person-person conversation is it? Is it the conversation between two people? Or just the students make a speech. I think the experts will also provide very professional suggestions about the questions types. (teacher)

Speaking of the detailed question types, my participants thought the specialists and the experts could refer to the opinions from many aspects. For example, from their years of learning, the students have formed their opinions about the English assessment system. S talked about the question types he thought could be used to assess Chinese students' English oral ability. He said,

For example, the first part is the students' self-introduction. It is a simple communication in people's daily life so that I don't think there are any problems in this part. The second part is the students' presentation about a certain topic. The presentation could be an introduction of an event or his/her

opinion about this event. In the third part, the students are given an article to read. Then, they are asked to summarize the article. In the last part, the students are asked to listen to a material first. Then, according to their understanding, they make a discussion. Like this, I think it is a more comprehensive oral test. (student)

According to the participants, the members of the professional linguistic education institutions need to do the research and development to generate a new English assessment model, and they need to undergo a process of fumbling and exploration.

Second, in addition to generating the new model of oral test, the marking system is also important. Participants talked about the challenges to be faced with when the model of person-person conversation is implemented. One of challenges is the level of the examining teachers' English oral proficiency. Another challenge is the examiners' total authority to mark the students' performances and the risk of bribes to give specific children good scores. These two situations put a negative influence on the marking process of the model of person-person conversation. Since the marking system is a necessary part of the English oral test, the linguistic specialists and the experts of language institutions have the responsibility to deal with these problems.

To solve this issue, R said,

One speech recording could be evaluated by two or three teachers. In this way, it is possible to handle the standards of marking. For example, if the marks given by two teachers are very different, there will be a third teacher to remark it or the leader of the marking group will remark it by himself or herself. Thus,



I think it is fair to some extent. (teacher)

Third, teachers training, qualification, and certifying is an important work the specialists and the experts need to deal with. When I asked the participants about the possibility to implement the model of person-person conversation in English oral test, most of them expressed concerns about the examiners' English proficiency, and about the inequity of English-language education caused by unqualified English teachers in remote areas of China. In this situation, they all emphasized the importance of teacher training. R (teacher) mentioned that one of the responsibilities of the specialists and the experts of language institutions is to train teachers in order to improve the English teachers' overall quality.

### **Suggestions to the First-Line English Teachers**

For the final level, the front-line English teachers, the participants also had several suggestions for implementing the model of person-person conversation. Since the English teachers are under the pressure of the higher education entry examination, they do not have much choice in class. Even if they understand the importance of English communicative capacities, they have to focus on the test-taking skills to make sure their students could pass the final higher education entry examination. However, if the State Educational Commission issues the policy to implement an oral English test in the higher education entry examination, the front-line teachers can work out ways to improve the students' oral ability and help them to get satisfactory scores in the test.

The 2 Chinese English teachers who were study participants had both worked

in a language-training center that focuses on IELTS training. Therefore, they were quite familiar with the pedagogy of teaching English oral classes. R mentioned,

At the beginning of the course, I give the students a test in order to know their real situation, to know at which English level they are. I give them some topics to talk about. At the very beginning, Chinese students are very shy, and they are afraid of making mistakes, so they are not willing to speak in English. Besides, they also have problems in making sentences. However, during the training, they have more opportunities to interact with the teacher, and they become relieved and are willing to speak in English. Then, at the final stage, I also give some topics in classes or have some seminars to ask the students to discuss in English. Hence, you will find the students making a progress in their communicative skills gradually. (teacher)

Q also shared her teaching experience. She said,

In my class, I try to talk to them in English. If they run across some new vocabulary they don't know, I will explain it in English first and give them some example sentences. I only tell them the Chinese meaning of the new vocabulary when it is really hard for the students to understand. In this way, I think it will be helpful for the students to think in English. (teacher)

Both of these teachers indicated that they teach their classes in English and that they communicate with their students as much as possible. They thought it was a good way to help the students to get used to the language environment and to think in English.

Although the teacher and the student participants gave suggestions for how to

achieve the model of person-person conversation, most of these suggestions were based on systemic changes. First, the participants thought that the State Educational Commission should issue the order to include the English oral test into the higher education entry examination first, but they all expressed doubts about this recommendation. No one knows when or whether the Commission would issue the order. Like H said, “it is an ideal model, but it’s not practical” (student).

### **Summary**

In this chapter, I presented the findings of the research. I analyzed the participants’ opinions about the English assessment system to evaluate the Chinese students’ communicative capacities. I organized their opinions into four different themes: (a) paper-written examination, (b) person-machine conversation, (c) person-person conversation, and (d) systemic changes for the model of person-person conversation.

The findings revealed the participants’ opinions about three different assessment models. Results showed that the most possible ways to assess Chinese students’ English communicative capacities were paper-written examination and person-machine conversation, although measures should be taken to improve these two models. On the other hand, the model of person-person conversation was the ideal assessment tool but was hard to achieve at the current stage.

## **CHAPTER FIVE: SUMMARY, DISCUSSION, AND IMPLICATIONS**

With more and more interaction between China and English-speaking countries, English communicative capacities have become much more important in China than before. In this situation, people wonder whether the traditional English assessment system could evaluate Chinese students' English oral skills effectively. To answer this question, I conducted a study to ask students' and teachers' opinions about how to develop a new assessment system for English communicative capacities in China. The findings of this study were intended to provide useful information to generate an effective examination model and help to improve Chinese students' English communicative capacities. In this chapter, I summarize the process of the study, the methodology, and the findings. Following the summary, I discuss the findings. Finally, I outline the implications for policy and practice.

### **Summary of the Study**

The study used qualitative research methodology. Individual telephone interviews were conducted with 2 Chinese English teachers and 2 Chinese students. I asked for their English teaching/learning experience in China and for their opinions about how to improve the current English assessment system to evaluate Chinese students' English communicative capacities effectively. The open-ended questions guided the process of data collection and data analysis.

Participants were drawn from one English Training Center in Shanghai, which specified in English language proficiency examination in order to prepare the students for going abroad. This language center offers English oral classes to help the students

pass different English oral tests and adapt to the life of English-speaking countries.

The participants had experienced both the traditional English examination in China and new English-teaching methods from English-speaking countries.

The interviews, which were conducted in Chinese, were audiotaped, transcribed, and translated into English. Data analysis revealed four major themes: (a) opinions about current paper-written examination, (b) opinions about the model of person-machine conversation, (c) opinions about the model of person-person conversation, and (d) systemic changes required for the model of person-person conversation.

Data analysis revealed that in both the teacher and the student participants' opinions, the ideal model was person-person conversation. However, it was not considered to be practical and realistic at the current stage. The most realistic ways to assess Chinese students' English communicative capacities were the paper-written examination and the model of person-machine conversation, but measures should be taken to improve these two models.

## **Discussion**

According to the major themes presented above, the discussion will be divided into three parts. The first part discusses the reasons why the model of person-person conversation could not be implemented at the current stage of China. Next, the necessity of paper-written examination and the revision of the question types are analyzed. Last, the importance of the model of person-machine conversation is outlined.

### **Challenges of the Model of Person-Person Conversation**

Data revealed that the model of person-person conversation is the ideal way to evaluate the students' English communicative capacities. However, the implementation involves a lot of systemic changes; therefore it is hard to put it into practice right away. During the interviews, both the teacher and the student participants preferred the exam model of person-person conversation. They thought that neither the paper-written examination nor the person-machine conversation could reproduce a real conversation situation. Qian (2009) also argues that the biggest difference between the model of person-person conversation and other models of English oral examination is the other models' "inability for the examiner and examinee to interact during the test, which appears to have created a psychological barrier for the test taker" (p. 113). As a result, person-person conversation appears to be the ideal model to assess a student's English communicative capacities. However, although the participants agreed that the model of person-person conversation was an effective assessment method, they identified three challenges that have to be dealt with.

The first challenge the participants pointed out is the quantity and the quality of English teachers in China. Specifically, this method would need 40,000 to 50,000 English teachers to be the examiners, and it is unclear whether these teachers would be qualified to be effective examiners. On this point, Rao (1996) has found that some Chinese English teachers lack English-language proficiency themselves. His findings confirm that it will be hard to find so many qualified English examiners because the

teachers do not have related training and there is no formal certificate to prove that an English teacher is qualified to become an examiner of English oral tests.

The second challenge is financial shortage to realize the model of person-person conversation. The financial shortage here refers to the economic difficulties that the government has to deal with. To implement the model of person-person conversation, the participants pointed out the need for enough English teachers to be the examiners to conduct the one-on-one oral tests for the higher education entry examination. Given the large amount of the senior-high graduates every year, the cost would be considerable.

The third challenge is equity. Hu (2012) says, “When deviations occur in terms of opportunity or quality, distortions may result, which cause educational inequities” (p. 3). If the model of person-person conversation is implemented, children from poor families might find it difficult to compete with their counterparts whose families can afford additional one-on-one English oral classes. In this case, the equity problem of educational resources would be caused by different family backgrounds. An equity problem also exists between urban areas and rural areas. One participant noted that if the model of person-person conversation was included in the higher education entry examination, the higher education would be out of reach for the students of rural areas. Feng (2009), who has investigated differences between rural and urban education, notes that rural primary schools did not start teaching English from Year Three until 2003. By contrast,

Learners of English in major economic and political powerhouses such as

Shanghai and Guangzhou may have access not only to teaching by qualified teachers at any age from kindergarten to an advanced level, but also to all modern facilities such as multimedia laboratories and computers, to private tuition by native speakers of English and even to tours abroad to gain direct exposure to English used in naturalistic settings. (p. 92)

Feng's study confirms that implementing the model of person-person conversation in the higher education entry examination system would increase the gap between the poor and the rich because of inequitable educational resources.

These challenges indicate how difficult it would be to conduct person-person conversation in the higher education entry examination at the current stage, although it appears to be the most effective way to evaluate students' English communicative capacities.

### **Improvement of Current Paper-Written Examination**

The paper-written examination is the most popular type in China. Although the participants doubted its validity to assess students' English communicative capacities, they agreed about its practical applicability in China at the current stage. One of the most important reasons is that the questions on paper-written examination are multiple-choice questions that can be easily scored by a machine. Their opinion is confirmed by Liu (2007), who claims that the multiple-choice questions can be easily administered in the classroom in a paper-and-pencil format that is easy to score. As a result, the paper-written examination is used as the primary assessment model of the higher education entry examination system.



However, the participants pointed out drawbacks to using the multiple-choice questions to evaluate the students' English communicative capacities, and their opinion is supported by the research literature. For example, the Learning Requirement of English Level at Primary and Secondary Stage in China requires that senior-high school graduates should "communicate fairly naturally with other English speakers about familiar topics" (Martin, 2005, p.7). However, this is difficult to assess by multiple-choice questions because, as Bacon (2003) points out, "multiple-choice questions are too simple a format to assess complex levels of knowledge or the ability to perform certain tasks" (p. 32). As an example, one teacher participant observed that the answers of multiple-choice questions have already been determined, but there are no such ready-made answers for daily conversations. In terms of this point, Rotfeld(1998) says, "Students proudly show off their high grades, from multiple-choice exams, as if their future careers will depend on knowing which choice to make instead of discerning which choices exist" (as cited in Bacon, 2003, p. 31). To overcome this drawback, the participants came up with the idea that new question types, such as short-answer questions, should be developed to assess students' English communicative capacities more effectively.

Short-answer questions are "tasks whose expected responses require a limited amount of language production" (Carr & Xi, 2010, p. 205). Compared with multiple-choice questions, short-answer questions have several advantages. First, students have to generate their own answers. About this point, Pressley et al. (1990) found that when students answer short-answer questions, they must construct their

own answers and do more processing than is necessary merely to discriminate between alternatives. Thus, short-answers questions imitate conversations in people's real lives better than multiple-choice questions.

Second, short-answer questions assess more aspects of students' English-language proficiency than multiple-choice. Regarding this point, Bensoussan and Kreindler (1990) claim that short-answer questions assess the students not only linguistically but also ideationally. For Chinese students whose first language is not English, it is difficult for them to understand each word of the conversation with people from English-speaking countries. However, for the aim of communication, it is not necessary to understand every word. If students can understand the content of the conversation, then they could still answer the questions in their own words. Besides, short-answer questions could also assess the students' spelling and vocabulary ability, which multiple-choice questions cannot do. In short, as a way to assess students' English communicative capacities, short-answer questions are more effective than multiple-choice questions.

Third, short-answer questions only require the students to use limited words to answer the questions, which makes it different from an English writing test. In oral communication, it is very possible for people to use incomplete sentences, which is not allowed in English writing tests, especially for academic essays. Given this aspect, short-answer questions could use this feature of oral language to assess the students' English communicative capacities, and the testing points of an oral communication test will not be mixed with the testing points of a writing test. Moreover, short-answer

questions could adjust the difficulty level of the answer format in addition to the content. For example, students at the senior level could be asked to answer a question in one or two complete sentences, but students at the junior level could be asked to answer the same question in a few words.

Although short-answer questions have many advantages, people still have concerns. The biggest concern is how to check the answers of short-answer questions because they do not have standardized answers. For multiple-choice questions, the answers could be checked by machines, which is the main reason why they are the most popular question types in China. In this way, it could save money and human resources. Although short-answer questions could be easily administered in classrooms, participants wondered whether teachers are able to score them.

In terms of scoring short-answer questions, an automated scoring system has been developed to help people assess the answers. Leacock (2004), for example, presented a system named C-rater, which “is intended for shorter answers and works by matching paraphrases of the concepts in correct answers” (as cited in Carr & Xi, 2010, p. 207). Although improvements still need to be made to the scoring system, as a question type that could assess students’ English communicative capacities, short-answer questions are effective and practical in China at the current stage.

### **Experimental Model of Person-Machine Conversation**

According to the participants and in the current situation, person-person conversation is not realistic to be put into practice, and there are also drawbacks to evaluate students’ English communicative capacities by paper-written examination.

Consequently, some advanced areas of China have implemented the English oral test of person-machine conversation as an experimental assessment model for the higher education entry examination. During the interviews, the participants expressed their opinions about this model. On one hand, they thought it could be the transition model between paper-written examination and person-person conversation. On the other hand, participants noted that the universities do not attach importance to it, which results in English communicative capacities still being ignored in the higher education entry examination system.

Since the model of person-person conversation is not practical at the current stage, the next choice is the model of person-machine conversation, which costs less. The policy of different exam models of the higher education entry examination in different places makes its implementation possible. In terms of this issue, Feng (2009) claims, “Even though China is a country that has been traditionally characterized by a centralized, top-down system, regional variations do exist in terms of state policy implementation” (p. 92). Some advanced areas, like Shanghai and Guangzhou, have implemented person-machine conversation first.

With regard to the English oral test of person-machine conversation, one of the student participants explained that students could be asked to answer the questions or make a presentation according to taped questions. Their performance would be recorded by computers and evaluated by more than one English teacher afterwards to make the process of assessment as fair as possible. For the advantages of this model, Qian (2009) says, “it is efficient, cost-effective, fair, and reliable” (p. 116). In fact,

many internationally oriented English-language proficiency tests have already used the model of person-machine conversation, like English for Speakers of Other Languages Examination (ESOL) and TOFEL.

Both teacher and student participants held positive opinions about the model of person-machine conversation. However, their biggest concern was that this model could not get the attention of enough people in China. In Shanghai, for example, for those students who applied for non-English-language related majors, the score of person-machine conversation was not counted into the total scores of the higher education entry examination. In Guangdong province, the marks of English oral test only make up a very small section of the total marks.

The higher education entry examination is the most important examination in China. Davey et al. (2007) have demonstrated its significance for individuals, families, and even the nation. As a result, in China, almost every teacher, student, and parent is exam-oriented. Both teacher and student participants noted that if the score of person-machine conversation was not included in or only made up a small section of the total scores, no one would attach importance to it. In the end, it would become a test just for a small group of people.

Besides, except for the purpose of assessing students' English-communicative capacities, it is also significant for the examination to improve students' English language proficiency. If it is just a test for the minority of people, it is difficult to achieve this goal in the whole nation. As a result, it is necessary to include the score of English oral tests of person-machine conversation into the total scores of the higher

education entry examination and make it important to arouse people's attention.

### **Implications**

This study provided insights into the development of the English assessment system in China. To better assess Chinese students' English communicative capacities, improvements should be made to the paper-written examination and the examination model of person-machine conversation. There are implications for both the policy level and the practice level, and for the three educational agencies of the State Education Commission as the superordinate, linguistics experts and specialists in language instruction as the intermediate, and classroom teachers as the subordinate agency (Adamson & Morris, 1997).

#### **Implications for Policies**

Policy changes are required to improve both the paper-written examination and the model of person-machine conversation. The State Educational Commission, the superordinate level, is the decision maker of educational policy in China (Adamson & Morris, 1997). It plays the core role in the development of Chinese English curriculum. For current paper-written examination, it needs to develop new short-answer questions. For this point, the State Educational Commission should issue orders to include short-answer questions into the oral part of the English test to make up for the disadvantages of multiple-choice questions.

For the experimental model of person-machine conversation in some urban areas, there are two reasons why it does not arouse enough attention there. First, in the higher education entry examination, only those senior-high graduates who apply for

majors related to English language are required to provide the scores of English oral test; therefore, for the majority who do not apply for those majors, they do not attach importance to it. Second, even for those graduates who are required to provide the scores of English oral test, the scores only make up a small portion of the total score of the higher education entry examination, which cannot arouse enough attention. To deal with these two problems, the State Educational Commission should issue orders to require all the universities to consider the applicants' scores of English oral test as one of the important admission requirements, no matter which majors those senior-high graduates apply for. In this way, English oral tests will play an important role in the higher education entry examination, and every teacher, student, and parent will not ignore it.

Because of the core role the higher education entry examination plays in the whole educational system, teachers, students, and parents are all exam-oriented. Only when the policy level makes the first change, will the other two agencies change accordingly.

### **Implications for Practice**

When the State Educational Commission makes the decision to innovate the paper-written examination and the model of person-machine conversation, the second and the third agencies will put the revolution into practice. The second agency, linguistics experts and specialists in language instruction, will be responsible for academic researches in order to instruct the revolution in details. According to Adamson and Morris (1997), linguistics experts and specialists are experienced and

professional English teachers. They are responsible for forming English assessments, writing questions of important examinations, training teachers, and so on. They are the people whom the State Educational Commission will turn to when the decision to develop the English assessment system is made.

For the development of short-answer questions, linguistics specialists and experts play an important role during the process. First, they can generate the detailed questions on the basis of academic researches that explore what kind of topics should be covered in this examination. For example, when Liu (2007) generated a multiple-choice discourse completion test (MDCT), he invited 30 Chinese university students to complete a questionnaire to help generate the multiple-choice questions. For the case of short-answer questions, as one part of the higher education entry examination, it is more important than MDCT. Thus, the complicated process of academic research should be handled by the specialists and experts.

Second, linguistics specialists and experts need to generate the marking standards of short-answer questions. Unlike multiple-choice questions, there are no standardized answers for short-answer questions. If there are no marking standards, even for the same answer, the score could vary from teacher to teacher. In terms of this point, Carr and Xi (2010) say, “Some errors that human raters make are random, such as those due to fatigue and lack of attention; others are systematic such as when some raters consistently misapply the scoring criteria” (p. 209). Currently, C-rater, an automated scoring system is available for short-answer questions. Carr and Xi believe that this system is better than human-raters. They explain, “Errors made by an



automated scoring engine, on the other hand, are systematic and predictable across all scored responses; random errors are not possible” (p. 209). By matching key terms or concepts, C-rater makes automated scoring possible. However, the specialists and experts are still needed to define the key terms and concepts. This process will continually take time and effort to make it complete and effective.

Third, linguistics specialists and experts are responsible for teacher training. For both paper-written examination and the model of person-machine conversation, if the portion of the content of English communicative capacities increases in the English test of the higher education entry examination, the concentration of English classes will be changed. To make sure the students could get satisfactory scores in the English test, English teachers will attach importance to English oral communicative capacities. However, according to Rao (1996), most Chinese English-language teachers are not prepared for this change. In this situation, professional training is necessary. As a professional linguistic educational agency, linguistics specialists and experts are responsible for teacher training to make them qualified.

As the third agency, English-language teachers carry out the innovation by directly teaching English communicative capacities to students. Compared with traditional English-language classes, the English classes that focus on communicative capacities will be different. New textbooks could be used, and pedagogy will be changed accordingly. The teachers are also in direct contact with the students. In this way, they could provide useful feedback about the new assessment system for the intermediate agency to consider.

In summary, for both paper-written examinations and the model of person-machine examinations, the innovation of English assessment will involve three educational agencies. Although the three agencies will need to collaborate with each other, the change of the intermediate and the subordinate agencies depends on the policy issued by the superordinate agency, the State Educational Commission. In the current exam-oriented situation, it is reasonable to have the top-down revolution process. Thus, the State Educational Commission should issue the related policy first. Then, the other two agencies will make a change accordingly.

### **Conclusion**

The purpose of this study was to develop a new assessment system to evaluate Chinese students' English communicative capacities. A qualitative research methodology was used to interview 4 participants for their opinions about this issue.

The findings of this study showed that although the participants thought the model of person-person conversation was the ideal way, it involved systemic changes, and it is impossible to realize at the current stage. They gave some suggestions to improve paper-written and person-machine conversation examinations in order to better evaluate students' English communicative capacities. According to the exam-oriented situation, a top-down change process should be conducted, and the three educational agencies that are involved need to cooperate with each other.

## References

- Adamson, B., & Morris, P. (1997). The English curriculum in the People's Republic of China. *Comparative Education Review*, 41(1), 3-26.
- Bacon, D. R. (2003). Assessing learning outcomes: A comparison of multiple-choice and short-answer questions in a marketing context. *Journal of Marketing Education*, 25(31), 31-36.
- Bensoussan, M., & Kreindler, I. (1990). Improving advanced reading comprehension in a foreign language: Summaries vs. short-answer questions. *Journal of Research in Reading*, 13(1), 55-68.
- Berg, B. L. (2001). *Qualitative research methods for the social sciences*. Boston, United States: Allyn and Bacon.
- Carr, N. T., & Xi, X. (2010). Automated scoring of short-answer reading items: Implications for constructs. *Language Assessment Quarterly*, 7, 205-218.
- Creswell, J. W. (2010). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Boston, United States: Pearson.
- Davey, G., Lian, C. D., & Higgins, L. (2007). The higher education entry examination system in china. *Journal of further and Higher Education*, 31(4), 385-396.
- Feng, A. (2009). English in China: Convergence and divergence in policy and practice. *AILA Review*, 22(1), 85-102.
- Hammond, B. G. (2010). A crack in the sorting machine: What we should not learn from China. *Independent School*, 69(3), 0.

- Hu, G. W. (2005). English language education in China: Policies, progress, and problems. *Language Policy*, 4(1), 5–24.
- Hu, W. (2012). Looking back at thirty years of educational reform and development in China. *Chinese Education and Society*, 45(1), 3-6.
- Li, C. (1990). An ESL motivations assessment for a community-based ESL program. *TESL Canada Journal*, 7(2), 31-42.
- Liu, J. (2007). Developing a pragmatics test for Chinese EFL learners. *Language Testing*, 24(3), 391-415.
- Li, X. (1984). In defense of the communicative approach. *ELT Journal*, 38(1), 2-13.
- Luo, J., & Wendel, F. C. (1999). Preparing for college: Senior high school education in china. *NASSP Bulletin*, 83(609), 57-68.
- Martin, T., (2005). New English curriculum for Chinese primary schools and junior/senior middle schools. Retrieved from <http://wenku.baidu.com/view/4426e428cfc789eb172dc82c.html>
- Mi, Z. H. (1998). The place of physics in Chinese middle schools. *Physics Teacher*, 23, 24.
- Pan, L. (2011). English language ideologies in the Chinese foreign language education policies: a world-system perspective. *Language Policy*, 10(3), 245-263.

- Pressley, M., Ghatala, E. S., Woloshyn, V., & Pirie, J. (1990). Sometimes adults miss the main ideas and do not realize it: Confidence in responses to short-answer and multiple-choice comprehension questions. *Reading Research Quarterly*, 25(3), 232-249.
- Qian, D. D. (2009). Comparing direct and semi-direct modes for speaking assessment: affective effects on test takers. *Language Assessment Quarterly*, 6, 113-125.
- Rao, Z. (1996). Reconciling communicative approaches to the teaching of English with traditional Chinese methods. *Research in the Teaching of English*, 30(4), 458-471.
- Song, X., & Cheng, L. (2006). Language learner strategy use and test performance of Chinese learners of English. *Lawrence Erlbaum Associates, Inc.*, 3(3), 243-266.
- Van Naerssen, M., & Riggenbach, H. (1987). Evaluating oral skills in an EST program: Native English-speaking scientists respond to Chinese scientists' English. *English for Specific Purposes*, 6(3), 219-232.
- Wang, W. F., & Lam, A. (2009). The English language curriculum for senior secondary school in China: Its evolution from 1949. *REL C Journal*, 40(1), 65–82.
- Zhang, C. (1995). The national higher education entry examination and its influence on secondary school physics teaching in China. *Physics Education*, 30, 104-108.

## Appendix A

### Interview Protocol for Teachers

Project: How to Develop a New English Assessment System for Communicative Capacities in China

Time of Interview:

Date:

Interviewer:

Interviewee:

[Describe here the project, reminding the interviewee about (a) the purpose of the study, (b) the individuals and sources of data being collected, (c) what will be done with the data to protect the confidentiality of the interviewee, (d) how long the interview will take, (e) the benefits the interviewees have, and (f) participants are free to withdraw at any time during this research and a written notice of that decision will be appreciated]

[Turn on the recorder and test it.]

Questions:

1. Please tell me your English teaching experience. (For example, how long have you been teaching English? What kind of English classes did you teach? What English skills you teach in these classes? Do these classes help your students improve their English communicative skills? How do you know that?)
  
2. Please share your opinions about the current English assessment system in China. (For example, what do you think about the English examination of the higher education entry examination system? Are there any parts of this examination related to English communicative skills? How effective do you think the higher education entry examination system is for evaluating students' English communicative skills? Why do you think this way? What connections and/or conflicts do you see between the assessment system for English communicative skills and the higher education entry examination system?)
  
3. Currently, we have two new models of assessment, one of which is to have the students' speech examples recorded, and to have teachers evaluate the recordings afterwards. Another one is multiple-choice questions. The questions are built around a conversation that supposedly happened in a certain circumstance, and

students are asked to choose the best response among the multiple choices. It has its own scoring system, with two points for the key, zero for the option that was incorrect in grammar or vocabulary, and one point for distractors, which are both linguistically inaccurate and socially inappropriate. What is your opinion of these two models of assessments? Do you think they can assess the students' English communicative capacities? Why or why not? Do you think they can be used effectively in China? Why or why not?

4. Thinking about the usual and the new models of English assessment, as well as your own experience, what would you say is the most effective assessment method to evaluate the students' English communicative skills? Why? (Feel free to bring in your own opinion.)
  
5. Thinking of your ideal English assessment system, do you think there are any conflicts between this new assessment system and higher education entry examination system? How could it be adapted it to the higher education entry examination system?
  
6. In China, there are three levels of agencies involved in the English curriculum development: (a) the State Education Commission as the superordinate, (b) linguistics experts and specialists in language instruction as the intermediate, and (c) the classroom teachers as the subordinate agency. What should these three agencies do to adopt your ideal English assessment system?
  
7. Are there any things you would like to add relevant to this topic?

(Thank the interviewees for their participation in this interview. Make confirmation that they will receive the transcript of the interview and the results of the summary.)

## **Appendix B**

### **Interview Protocol for Students**

Project: How to Develop a New English Assessment System for Communicative Capacities in China

Time of Interview:

Date:

Interviewer:

Interviewee:

[Describe here the project, reminding the interviewee about (a) the purpose of the study, (b) the individuals and sources of data being collected, (c) what will be done with the data to protect the confidentiality of the interviewee, (d) how long the interview will take, (e) the benefits the interviewees have, and (f) participants are free to withdraw at any time during this research and a written notice of that decision will be appreciated]

[Turn on the recorder and test it.]

Questions:

1. Please tell me your English learning experience. (For example, how long have you been studying English? What kind of English classes did you have? What English skills you learn from these classes? Do these classes help you improve your English communicative skills? How do you know that?)
2. Please share your opinions about the current English assessment system in China. (For example, what do you think about the English examination of the higher education entry examination system? Are there any parts of this examination related to English communicative skills? How effective do you think the higher education entry examination system is for evaluating students' English communicative skills? Why do you think this way? What connections and/or conflicts do you see between the assessment system for English communicative skills and the higher education entry examination system?)
3. Currently, we have two new models of assessment, one of which is to have the students' speech examples recorded, and to have teachers evaluate the recordings afterwards. Another one is multiple-choice questions. The questions are built around a conversation that supposedly happened in a certain circumstance, and



students are asked to choose the best response among the multiple choices. It has its own scoring system, with two points for the key, zero for the option that was incorrect in grammar or vocabulary, and one point for distractors, which are both linguistically inaccurate and socially inappropriate. What is your opinion of these two models of assessments? Do you think they can assess the students' English communicative capacities? Why or why not? Do you think they can be used effectively in China? Why or why not?

4. Thinking about the usual and the new models of English assessment, as well as your own experience, what would you say is the most effective assessment method to evaluate the students' English communicative skills? Why? (Feel free to bring in your own opinion.)
  
5. Thinking of your ideal English assessment system, do you think there are any conflicts between this new assessment system and higher education entry examination system? How could it be adapted it to the higher education entry examination system?
  
6. In China, there are three levels of agencies involved in the English curriculum development: (a) the State Education Commission as the superordinate, (b) linguistics experts and specialists in language instruction as the intermediate, and (c) the classroom teachers as the subordinate agency. What should these three agencies do to adopt your ideal English assessment system?
  
7. Are there any things you would like to add relevant to this topic?

(Thank the interviewees for their participation in this interview. Make confirmation that they will receive the transcript of the interview and the results of the summary.)